

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Patent Application of	)	
Yoko FUJIWARA	)	Group Art Unit: Unassigned
Application No.: Unassigned	)	Examiner: Unassigned
Filed: May 7, 2001	)	
For: IMAGE RECOGNITION	)	
APPARATUS...	)	
	)	
	)	
	)	
	)	

11040 U.S. PTO  
09/850094  
05/08/01

**CLAIM FOR CONVENTION PRIORITY**

Assistant Commissioner for Patents  
Washington, D.C. 20231

Sir:

The benefit of the filing date of the following prior foreign application in the following foreign country is hereby requested, and the right of priority provided in 35 U.S.C. § 119 is hereby claimed:

Japanese Patent Application No. 2000-136301

Filed: May 9, 2000

In support of this claim, enclosed is a certified copy of said prior foreign application. Said prior foreign application was referred to in the oath or declaration. Acknowledgment of receipt of the certified copy is requested.

Respectfully submitted,

BURNS, DOANE, SWECKER & MATHIS, L.L.P.

Date: May 8, 2001

By: William Choukeng RN 30888, for  
Platon N. Mandros  
Registration No. 22,124

P.O. Box 1404  
Alexandria, Virginia 22313-1404  
(703) 836-6620

日 本 国 特 許 庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

J11040 U.S. PTO  
09/850094  
05/08/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

2000年 5月 9日

出 願 番 号

Application Number:

特願2000-136301

出 願 人

Applicant(s):

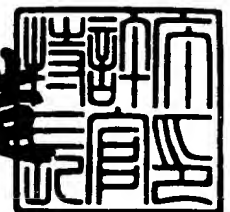
ミノルタ株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 2月 2日

特許庁長官  
Commissioner,  
Patent Office

及川耕造



出証番号 出証特2001-3004534

【書類名】 特許願

【整理番号】 AK05204

【提出日】 平成12年 5月 9日

【あて先】 特許庁長官 近藤 隆彦 殿

【国際特許分類】 G06K 9/20

【発明の名称】 画像認識装置、画像認識方法および画像認識プログラム  
を記録したコンピュータ読取可能な記録媒体

【請求項の数】 10

【発明者】

【住所又は居所】 大阪府中央区安土町二丁目3番13号 大阪国際ビル  
ミノルタ株式会社内

【氏名】 藤原 葉子

【特許出願人】

【識別番号】 000006079

【氏名又は名称】 ミノルタ株式会社

【代理人】

【識別番号】 100072349

【弁理士】

【氏名又は名称】 八田 幹雄

【電話番号】 03-3230-4766

【選任した代理人】

【識別番号】 100102912

【弁理士】

【氏名又は名称】 野上 敦

【選任した代理人】

【識別番号】 100110995

【弁理士】

【氏名又は名称】 奈良 泰男

【選任した代理人】

【識別番号】 100111464

【弁理士】

【氏名又は名称】 齋藤 悦子

【選任した代理人】

【識別番号】 100114649

【弁理士】

【氏名又は名称】 宇谷 勝幸

【手数料の表示】

【予納台帳番号】 001719

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 画像認識装置、画像認識方法および画像認識プログラム  
を記録したコンピュータ読取可能な記録媒体

【特許請求の範囲】

【請求項 1】 画像認識装置において、

入力された画像データの中から文字画像が存在する文字領域を検出する文字領域検出部と、

文字画像から文字コードを認識して文字コードデータを得る文字認識部と、

得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する文字領域外判別部と、

前記入力された画像データ中の文字領域に隣接し、かつ、文字画像、または非文字である図形画像のいずれをも含まない所定量の未使用部分が存在するか否かを判別する未使用部分判別部と、

前記文字領域外判別部により、前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合であって、前記未使用部分判別部により、前記未使用部分が存在すると判別されたとき、文字コードデータが配置される前記文字領域を、前記未使用部分を含めた領域まで拡大する文字領域変更部と、

を有することを特徴とする画像認識装置。

【請求項 2】 前記文字領域外判別部により、前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合であって、前記未使用部分判別部により、前記未使用部分が存在しないと判別されたとき、当該文字コードデータの文字サイズを小さくする文字サイズ変更部を、さらに有することを特徴とする請求項 1 に記載の画像認識装置。

【請求項 3】 文字領域内で文字画像が複数の連続しない独立した列を形成しているか否かを判別する列判別部を、さらに有し、

前記列判別部により、前記文字領域内の各列が独立した列であると判別された場合、前記文字領域外判別部および前記未使用部分判別部は、前記入力された画

像データ中の文字領域に代えて、当該文字領域を列単位で分割した個々の列領域を使用して判別することを特徴とする請求項 1 または 2 に記載の画像認識装置。

【請求項 4】 前記列判別部は、複数列を形成する文字画像の各列についての、隣接する列に対する始点位置および終点位置の変化量が所定値よりも大きい場合、文字画像が複数の連続しない独立した列を形成していると判別することを特徴とする請求項 3 に記載の画像認識装置。

【請求項 5】 画像認識装置において、

入力された画像データの中から文字画像が存在する文字領域を検出する文字領域検出部と、

前記文字領域を含み当該文字領域の背景となる背景領域であって、他の異なった背景領域の内部に存在する文字背景領域を検出する文字背景領域検出部と、

文字画像から文字コードを認識して文字コードデータを得る文字認識部と、

得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する文字領域外判別部と、

前記文字領域外判別部により、前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合、文字コードデータが配置される前記文字領域を前記文字背景領域の範囲内で拡大する文字領域変更部と、

前記文字領域変更部により、前記文字領域を前記文字背景領域と一致するまで拡大した場合であって、得られた文字コードデータを所定のフォントで配置したとき、前記文字背景領域から文字コードデータがはみだすか否かを判別する背景領域外判別部と、

前記背景領域外判別部により、前記文字背景領域から文字コードデータがはみだすと判別された場合、当該文字コードデータの文字サイズを小さくする文字サイズ変更部と、

を有することを特徴とする画像認識装置。

【請求項 6】 文字領域内で文字画像が複数の連続しない独立した列を形成しているか否かを判別する列判別部と、

前記列判別部により、前記文字領域内の各列が独立した列であると判別された

場合、当該各列が、前記文字背景領域内において左寄せ、右寄せ、またはセンタリングのいずれの形式で配置されているかを検出する文字配置検出部と、

前記文字配置検出部の検出結果に基づいて、前記各列の配置形式を指定する配置形式指定部と、をさらに有することを特徴とする請求項 5 に記載の画像認識装置。

【請求項 7】 画像認識方法において、

入力された画像データの中から文字画像が存在する文字領域を検出する過程と

、  
文字画像から文字コードを認識して文字コードデータを得る過程と、

得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する過程と、

前記入力された画像データ中の文字領域に隣接し、かつ、文字画像、または非文字画像である図形画像のいずれをも含まない所定量の未使用部分が存在するか否かを判別する過程と、

前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合であって、前記未使用部分が存在すると判別されたとき、文字コードデータが配置される前記文字領域を前記未使用部分を含めた領域まで拡大する過程と、

を有することを特徴とする画像認識方法。

【請求項 8】 画像認識方法において、

入力された画像データの中から文字画像が存在する文字領域を検出する過程と

、  
前記文字領域を含み当該文字領域の背景となる背景領域であって、他の異なった背景領域の内部に存在する文字背景領域を検出する過程と、

文字画像から文字コードを認識して文字コードデータを得る過程と、

得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する過程と、

前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合、文字コードデータが配置される前記文字領域を前記文字背景領域の範囲内で拡大する過程と、

前記文字領域を前記文字背景領域と一致するまで拡大した場合であって、得られた文字コードデータを所定のフォントで配置したとき、前記文字背景領域から文字コードデータがはみだすか否かを判別する過程と、

前記文字背景領域から文字コードデータがはみだすと判別された場合、当該文字コードデータの文字サイズを小さくする過程と、

を有することを特徴とする画像認識方法。

【請求項 9】 画像認識プログラムを記録したコンピュータ読取可能な記録媒体であって、

入力された画像データの中から文字画像が存在する文字領域を検出する過程と、

文字画像から文字コードを認識して文字コードデータを得る過程と、

得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する過程と、

前記入力された画像データ中の文字領域に隣接し、かつ、文字画像、または非文字画像である図形画像のいずれをも含まない所定量の未使用部分が存在するか否かを判別する過程と、

前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合であって、前記未使用部分が存在すると判別されたとき、文字コードデータが配置される前記文字領域を前記未使用部分を含めた領域まで拡大する過程と、

をコンピュータに実行させるための画像認識プログラムを記録したコンピュータ読取可能な記録媒体。

【請求項 10】 画像認識プログラムを記録したコンピュータ読取可能な記録媒体であって、

入力された画像データの中から文字画像が存在する文字領域を検出する過程と



前記文字領域を含み当該文字領域の背景となる背景領域であって、他の異なった背景領域の内部に存在する文字背景領域を検出する過程と、

文字画像から文字コードを認識して文字コードデータを得る過程と、

得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する過程と、

前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合、文字コードデータが配置される前記文字領域を前記文字背景領域の範囲内で拡大する過程と、

前記文字領域を前記文字背景領域と一致するまで拡大した場合であって、得られた文字コードデータを所定のフォントで配置したとき、前記文字背景領域から文字コードデータがはみだすか否かを判別する過程と、

前記文字背景領域から文字コードデータがはみだすと判別された場合、当該文字コードデータの文字サイズを小さくする過程と、

をコンピュータに実行させるための画像認識プログラムを記録したコンピュータ読取可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、画像認識技術に関し、特に、入力された画像データ中の文字画像から文字コードを認識して得られた文字コードデータを、ある特定のフォントで配置した場合でも、元のレイアウトを維持する機能を備えた画像認識技術に関する。

【0002】

【従来の技術】

画像認識装置は、例えばスキャナ等で原稿を読み取ることにより入力された画像データを解析して、画像データの属性ごとに、文字画像、写真画像、あるいは図表画像が存在する各領域を抽出し、各領域についてそれぞれの属性に応じた認

識処理を行う。

【0003】

最近では、単に文字画像から文字コードを認識して得られた文字コードデータ  
を出力するだけではなく、元の原稿のレイアウトを維持して出力する機能を備え  
た画像認識装置が提案されてきている。

【0004】

例えば、文字画像から、フォントサイズや、フォントのタイプ、スタイル（太  
字、斜体等）などの文字属性を認識し、これらの文字属性をも含めて文字コード  
データを出力することにより、原稿レイアウトの維持を図った画像認識装置が知  
られている。

【0005】

また、写真画像や図表画像をビットマップデータとして埋め込んだり、段組み  
を認識してレイアウト枠の形式で記述したりして、各種形式のファイルを作成す  
る画像認識装置も知られている。

【0006】

【発明が解決しようとする課題】

ところが、上記した従来の画像認識装置は、文字画像から、文字属性としてあ  
る程度のフォントのタイプを認識することができるに過ぎない。しかも、例えば  
原稿上に存在するフォントの種類は非常に多いことから、すべてのフォントを識  
別することも現実的ではない。また仮に、原稿上のフォントをすべて正確に識別  
することができ、識別されたフォントを指定して出力したとしても、例えばデー  
タを受け取る側に当該フォントが用意されていなければ、結局のところ正確なフ  
ォントでデータを受け取ることはできない。

【0007】

このため、現実的には、例えば原稿上に存在する数多くの種類のフォント（以  
下、「入力フォント」ともいう。）は、画像認識装置が文字画像から文字コード  
を認識して文字コードデータで出力する際に使用する一般的な数種類のフォント  
（以下、「出力フォント」ともいう。）に対応させられることになる。

【0008】

したがって、フォントの種類によって文字幅や文字高さが異なることを考えれば、従来の画像認識装置は、例えば原稿を読み取ることによって入力された画像データ中の文字画像から文字コードを認識して得られた文字コードデータを特定の出力フォントで出力した場合、原稿における文字の配置と同じ配置で文字コードデータを出力することができない場合が多くなる。

## 【 0 0 0 9 】

結果的に、従来の画像認識装置は、原稿レイアウトの維持機能を備えたものであっても、實際上、種々の原稿に対応することが困難である。具体的には、原稿上のフォント、あるいは文字や図形等のレイアウトによっては、原稿を読み取ることによって入力された文字画像から文字コードを認識して得られた文字コードデータを特定の出力フォントで例えば用紙に印刷した場合、以下のような問題があった。

## 【 0 0 1 0 】

例えば出力フォントの文字幅の方が入力フォントの文字幅よりも大きい場合、原稿を読み取ることにより入力された画像データ中における文字画像が存在する文字領域内に、出力フォントの文字コードデータに変換された文字画像が収まらなくなってしまう事態が発生する。このため、印刷された用紙上で例えば文字と写真とが重なったり、あるいは文字同士が重なったりするという問題があった。さらには、文字列の数が増加してしまっ、て、1 頁の原稿を読み取ったにも拘らず、2 頁の用紙に印刷されてしまうという問題もあった。一方、用紙上での例えば文字と写真との重なりを避けるために、両者を離間させる処理が実行されると、写真だけが別の頁の用紙に印刷されてしまうこともある。

## 【 0 0 1 1 】

そこで、本発明の目的は、入力された画像データ中の文字画像から文字コードを認識して得られた文字コードデータをある特定のフォントで配置した場合でも、元のレイアウトを適切に維持しつつ、しかも、文字コードデータを他の文字コードデータや図形画像と重なることなく同一頁内に配置することにある。

## 【 0 0 1 2 】

【課題を解決するための手段】

本発明の目的は、下記する手段により達成される。

【0013】

(1) 画像認識装置において、入力された画像データの中から文字画像が存在する文字領域を検出する文字領域検出部と、文字画像から文字コードを認識して文字コードデータを得る文字認識部と、得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する文字領域外判別部と、前記入力された画像データ中の文字領域に隣接し、かつ、文字画像、または非文字である図形画像のいずれをも含まない所定量の未使用部分が存在するか否かを判別する未使用部分判別部と、前記文字領域外判別部により、前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合であって、前記未使用部分判別部により、前記未使用部分が存在すると判別されたとき、文字コードデータが配置される前記文字領域を、前記未使用部分を含めた領域まで拡大する文字領域変更部と、を有することを特徴とする画像認識装置。

【0014】

(2) 前記文字領域外判別部により、前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合であって、前記未使用部分判別部により、前記未使用部分が存在しないと判別されたとき、当該文字コードデータの文字サイズを小さくする文字サイズ変更部を、さらに有することを特徴とする上記(1)に記載の画像認識装置。

【0015】

(3) 文字領域内で文字画像が複数の連続しない独立した列を形成しているか否かを判別する列判別部を、さらに有し、前記列判別部により、前記文字領域内の各列が独立した列であると判別された場合、前記文字領域外判別部および前記未使用部分判別部は、前記入力された画像データ中の文字領域に代えて、当該文字領域を列単位で分割した個々の列領域を使用して判別することを特徴とする上記(1)または(2)に記載の画像認識装置。

【0016】

(4) 前記列判別部は、複数列を形成する文字画像の各列についての、隣接

する列に対する始点位置および終点位置の変化量が所定値よりも大きい場合、文字画像が複数の連続しない独立した列を形成していると判別することを特徴とする上記（３）に記載の画像認識装置。

## 【 0 0 1 7 】

（５） 画像認識装置において、入力された画像データの中から文字画像が存在する文字領域を検出する文字領域検出部と、前記文字領域を含み当該文字領域の背景となる背景領域であって、他の異なった背景領域の内部に存在する文字背景領域を検出する文字背景領域検出部と、文字画像から文字コードを認識して文字コードデータを得る文字認識部と、得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する文字領域外判別部と、前記文字領域外判別部により、前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合、文字コードデータが配置される前記文字領域を前記文字背景領域の範囲内で拡大する文字領域変更部と、前記文字領域変更部により、前記文字領域を前記文字背景領域と一致するまで拡大した場合であって、得られた文字コードデータを所定のフォントで配置したとき、前記文字背景領域から文字コードデータがはみだすか否かを判別する背景領域外判別部と、前記背景領域外判別部により、前記文字背景領域から文字コードデータがはみだすと判別された場合、当該文字コードデータの文字サイズを小さくする文字サイズ変更部と、を有することを特徴とする画像認識装置。

## 【 0 0 1 8 】

（６） 文字領域内で文字画像が複数の連続しない独立した列を形成しているか否かを判別する列判別部と、前記列判別部により、前記文字領域内の各列が独立した列であると判別された場合、当該各列が、前記文字背景領域内において左寄せ、右寄せ、またはセンタリングのいずれの形式で配置されているかを検出する文字配置検出部と、前記文字配置検出部の検出結果に基づいて、前記各列の配置形式を指定する配置形式指定部と、をさらに有することを特徴とする上記（５）に記載の画像認識装置。

## 【 0 0 1 9 】

(7) 画像認識方法において、入力された画像データの中から文字画像が存在する文字領域を検出する過程と、文字画像から文字コードを認識して文字コードデータを得る過程と、得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する過程と、前記入力された画像データ中の文字領域に隣接し、かつ、文字画像、または非文字画像である図形画像のいずれをも含まない所定量の未使用部分が存在するか否かを判別する過程と、前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合であって、前記未使用部分が存在すると判別されたとき、文字コードデータが配置される前記文字領域を前記未使用部分を含めた領域まで拡大する過程と、を有することを特徴とする画像認識方法。

## 【0020】

(8) 画像認識方法において、入力された画像データの中から文字画像が存在する文字領域を検出する過程と、前記文字領域を含み当該文字領域の背景となる背景領域であって、他の異なった背景領域の内部に存在する文字背景領域を検出する過程と、文字画像から文字コードを認識して文字コードデータを得る過程と、得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する過程と、前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合、文字コードデータが配置される前記文字領域を前記文字背景領域の範囲内で拡大する過程と、前記文字領域を前記文字背景領域と一致するまで拡大した場合であって、得られた文字コードデータを所定のフォントで配置したとき、前記文字背景領域から文字コードデータがはみだすか否かを判別する過程と、前記文字背景領域から文字コードデータがはみだすと判別された場合、当該文字コードデータの文字サイズを小さくする過程と、を有することを特徴とする画像認識方法。

## 【0021】

(9) 画像認識プログラムを記録したコンピュータ読取可能な記録媒体であって、入力された画像データの中から文字画像が存在する文字領域を検出する過

程と、文字画像から文字コードを認識して文字コードデータを得る過程と、得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する過程と、前記入力された画像データ中の文字領域に隣接し、かつ、文字画像、または非文字画像である図形画像のいずれをも含まない所定量の未使用部分が存在するか否かを判別する過程と、前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合であって、前記未使用部分が存在すると判別されたとき、文字コードデータが配置される前記文字領域を前記未使用部分を含めた領域まで拡大する過程と、をコンピュータに実行させるための画像認識プログラムを記録したコンピュータ読取可能な記録媒体。

## 【 0 0 2 2 】

(10) 画像認識プログラムを記録したコンピュータ読取可能な記録媒体であって、入力された画像データの中から文字画像が存在する文字領域を検出する過程と、前記文字領域を含み当該文字領域の背景となる背景領域であって、他の異なった背景領域の内部に存在する文字背景領域を検出する過程と、文字画像から文字コードを認識して文字コードデータを得る過程と、得られた文字コードデータを所定のフォントで配置した場合、前記入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する過程と、前記入力された画像データ中の文字領域から文字コードデータがはみだすと判別された場合、文字コードデータが配置される前記文字領域を前記文字背景領域の範囲内で拡大する過程と、前記文字領域を前記文字背景領域と一致するまで拡大した場合であって、得られた文字コードデータを所定のフォントで配置したとき、前記文字背景領域から文字コードデータがはみだすか否かを判別する過程と、前記文字背景領域から文字コードデータがはみだすと判別された場合、当該文字コードデータの文字サイズを小さくする過程と、をコンピュータに実行させるための画像認識プログラムを記録したコンピュータ読取可能な記録媒体。

## 【 0 0 2 3 】

## 【発明の実施の形態】

以下、添付した図面を参照して、本発明の実施の形態を説明する。

## 【 0 0 2 4 】

図 1 は、本発明の画像認識装置の第 1 の実施形態に係るデジタルコピー機を含む情報機器システムの構成を示す概略ブロック図である。

## 【 0 0 2 5 】

図 1 に示すように、情報機器システム 1 0 0 は、デジタルコピー機 1 0、およびコンピュータ 5 0 等の情報機器がネットワーク N 上に接続されて構成されている。

## 【 0 0 2 6 】

画像認識装置としてのデジタルコピー機 1 0 は、デジタルコピー機の各部を統括制御する主制御部 1 1 と、原稿を読み取って画像データを得るスキャナ部 1 2 と、画像データを解析して、文字画像から文字コードを認識して文字コードデータを得たり、データ圧縮等の処理を行ったりする画像処理部 1 3 と、所定のデータを用紙に印刷するプリント部 1 4 と、あらかじめ作成された所定のプログラムやデータを記憶する ROM 1 5 と、一時的にデータを記憶する RAM 1 6 と、各種情報を表示したりユーザーが指示を与えたりするための操作パネル 1 7 と、コンピュータ 5 0 等の各情報機器との間でコマンドやデータの送受信を行うためのインタフェース ( I / F ) 1 8 とを有している。

## 【 0 0 2 7 】

デジタルコピー機 1 0 とコンピュータ 5 0 とは、ネットワーク N を経由して互いにデータの授受を行うことが可能となっている。例えば、デジタルコピー機 1 0 は、スキャナ部 1 2 で得られた画像データを画像処理部 1 3 で画像処理した上で、コンピュータ 5 0 に送信することができる。また、デジタルコピー機 1 0 は、コンピュータ 5 0 から送信された画像データを受信して、プリント部 1 4 により用紙に印刷することもできる。

## 【 0 0 2 8 】

すなわち、デジタルコピー機 1 0 は、スキャナ部 1 2 で原稿を読み取ることによって入力された画像データを必要に応じて画像処理した上でプリント部 1 4 で用紙に印刷するコピー機能のほか、スキャナ部 1 2 を独立して使用するスキャン機能や、プリント部 1 4 を独立して使用するプリント機能を有している。



## 【 0 0 2 9 】

スキャナ部 1 2 で原稿を読み取ることによって入力された画像データは、例えば、画像処理部 1 3 に送られ、画像処理部 1 3 は、画像データに含まれる文字画像から文字コードを認識して文字コードデータに変換した上で、所定の形式のファイルを作成する。作成されたファイルは、デジタルコピー機 1 0 の操作パネル 1 7 によりコンピュータ 5 0 への取り込みが指示された場合、コンピュータ 5 0 に送信され、操作パネル 1 7 によりコピーが指示された場合、プリント部 1 4 に送信される。また、作成されたファイルを、ネットワーク N を介して電子メールに添付して送信することも可能である。

## 【 0 0 3 0 】

図 1 に示すように、画像処理部 1 3 は、レイアウト解析部 3 1、文字認識部 3 2、フォント認識部 3 3、文字領域変更部 3 4、文字サイズ変更部 3 5、およびファイル作成部 3 6 を備えている。

## 【 0 0 3 1 】

レイアウト解析部 3 1 は、入力された画像データを解析し、画像データの属性ごとに、文字画像が存在する文字領域、非文字画像である図形画像が存在する写真領域、あるいはグラフやチャートのような、文字画像、および非文字画像である図形画像が混在する図表領域をそれぞれ抽出する。

## 【 0 0 3 2 】

文字領域内の文字画像は、後述するように、文字認識部 3 2 により文字コードが認識されて、文字コードデータに変換される。写真領域内の図形画像は、入力された画像データのままビットマップデータで、例えばファイルに保存される。なお、図表領域内の文字画像は、ユーザーによる操作パネル 1 7 上での設定の切り替えによって、文字コードデータに変換するか否かが決定される。

## 【 0 0 3 3 】

文字、写真および図表の各領域の抽出は、既知の方法により行われるが、例えば、文字領域の抽出は、画像データ中で図形的に連結している部分の大きさを測定し、文字画像の方がそれ以外の図形画像に比べて一般にサイズが小さいことを利用して行われる。なお、本実施形態では、レイアウト解析は、多値画像データ

を使用して行われるが、2 値画像データを使用して行ってもよい。

【 0 0 3 4 】

文字認識部 3 2 は、個々の文字画像から文字コードを認識して文字コードデータを得るほか、個々の文字画像の位置情報を認識する。各文字画像の位置情報は、例えば各文字画像の外接矩形の左上および右下の座標として得られる。文字画像から文字コードを認識する方法は、例えば、各文字画像の特徴量と ROM 1 5 に記憶されている辞書パターンとの一致の度合いに基づいて行われる。

【 0 0 3 5 】

フォント認識部 3 3 は、文字画像から、フォントの種類とフォントサイズとを認識する。フォントは、例えば上記した文字コードデータの認識方法と同様な方法により、所定の数種類の代表的な出力フォントに分類されて判別される。

【 0 0 3 6 】

文字領域変更部 3 4 は、文字コードデータが配置される文字領域を、原稿を読み取ることによって入力された画像データ中の文字領域に対して、広げる調整を行うことができる。また、文字サイズ変更部 3 5 は、文字コードデータの文字サイズを小さくする調整を行うことができる。

【 0 0 3 7 】

ファイル作成部 3 6 は、入力された画像データに対し、文字画像から文字コードを認識して文字コードデータに変換する等の所定の処理を施した上で、所定の形式のファイルを作成する。

【 0 0 3 8 】

次に、このように構成された画像認識装置としてのデジタルコピー機の作用について説明する。

【 0 0 3 9 】

図 2 は、画像認識処理を示すフローチャート、図 3 は、図 2 に示される出力調整処理を示すサブルーチンのフローチャートである。

【 0 0 4 0 】

あらかじめ、スキャナ部 1 2 により原稿を読み取ることによって画像データが入力される。入力された画像データは、多値画像データである。

## 【 0 0 4 1 】

まず、OCR前処理が行われる（S 1 1）。このOCR前処理では、入力された画像データ中のノイズの除去や、原稿が傾いて読み取られた場合の補正処理等が行われる。

## 【 0 0 4 2 】

次に、レイアウト解析が行われる（S 1 2）。図 4 は、原稿の一例を示す図、図 5 は、図 4 の原稿を読み取ることによって入力された画像データをレイアウト解析して文字領域を検出した結果を示す図である。図 5 に示すように、レイアウト解析部 3 1 により、文字画像が存在する矩形の文字領域 A ～ E が検出されて、抽出される。なお、図中の文字領域 A，B，C，E 内における横方向の実線は、個々の文字画像が並んで形成された横の列（以下、「行」という。）を示す。

## 【 0 0 4 3 】

レイアウト解析の後、文字領域におけるカラーまたはモノクロの多値画像データに対して 2 値化処理を行うことにより、モノクロ 2 値画像データを作成する（S 1 3）。これにより、モノクロ 2 値画像データで表現された文字画像が得られることになる。

## 【 0 0 4 4 】

次いで、文字認識部 3 2 により、個々の文字画像から文字コードが認識されて文字コードデータが得られる。また、個々の文字画像の位置情報が認識される（S 1 4）。さらに、フォント認識部 3 3 により、文字画像からフォントの種類とフォントサイズとが認識される（S 1 5）。ここで、フォントは、画像認識装置に用意されている数種類の代表的な出力フォントの中から認識される。

## 【 0 0 4 5 】

そして、得られた文字コードデータをフォント認識部 3 3 により認識された出力フォントで配置すると、原稿を読み取ることによって入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別し、この判別結果に基づいて、出力調整処理を行う（S 1 6）。なお、出力フォントとしては、通常、フォント認識部 3 3 により認識されたフォントが適用されるが、ユーザーが任意に設定し得るように構成することも可能である。

## 【 0 0 4 6 】

この出力調整処理では、図 3 に示すように、まず、文字領域内の行のピッチを算出する（S 2 1）。行ピッチは、文字領域の高さを行数で割ることにより算出される。ここで、文字領域の高さや幅は、レイアウト解析（S 1 2）の結果から得られる。また、行数は、個々の文字画像の位置情報としての外接矩形の x 座標（横方向の座標）を 1 文字画像ずつ順に調べていって、上記 x 座標がマイナス方向（逆方向）に変化した場合に改行があったと判断することにより、得ることができる。但し、行数が 1 行からなる文字領域については、行ピッチを算出することができない。この場合、行ピッチを設定しないという判断が行われる。

## 【 0 0 4 7 】

次いで、文字領域内の行が連続行であるか否かの判別を行う（S 2 2）。図 5 の文字領域 A は、文字領域内の行が連続行である一例を示しており、文字領域 B は、文字領域内の行が独立行である一例を示している。

## 【 0 0 4 8 】

図 6 は、文字領域内の行が連続行であるか否かの判別方法を説明するための図である。文字領域内の行が連続行であるか否かの判別は、各行の始点位置および終点位置の変化量を検出することにより行うことができる。具体的には、複数行を形成する文字画像の各行についての、隣接する行に対する始点位置および終点位置の変化量の平均値が、所定値よりも大きい場合、文字領域内の行が独立行であると判別し、所定値よりも小さい場合、文字領域内の行が連続行であると判別する。

## 【 0 0 4 9 】

すなわち、図 6 に示すように、 $S_i$  ( $i = 1, 2, \dots$ ) を  $i$  行目の始点 x 座標、 $E_i$  ( $i = 1, 2, \dots$ ) を  $i$  行目の終点 x 座標、 $N$  を行数として、

$$\Sigma \{ (S_{i+1} - S_i) + (E_{i+1} - E_i) \} / N$$

を計算し、当該値の大小により文字領域内の行が連続行であるか否かを判別する。但し、図 6 に示すように、読点やピリオド等の文の終りを示す文字画像を含む行  $L$  は、終点位置の変化量はないものとして計算する。

## 【 0 0 5 0 】

以上までのステップで、文字領域の高さと幅、行ピッチと行数、文字領域内の行が連続行であるか否かの判別結果、文字コード、フォントサイズ、フォントの種類、各文字画像の位置情報（外接矩形の座標）、および出力に用いるフォントメトリックデータを取得することができる。ここで、フォントメトリックデータとは、フォントの文字幅や文字高さ等の寸法データをいう。本実施形態では、フォントメトリックデータとして、文字幅を用いる。

## 【 0 0 5 1 】

文字領域内の行が連続行であれば、フォント認識部 3 3 により認識された出力フォントが、原稿上の文字の入力フォントと相違することにより多少文字幅が変化したとしても、原稿を読み取ることによって入力された画像データ中の文字領域に、文字画像から文字コードを認識して得られた文字コードデータが全体として配置することができればよい。一方、文字領域内の行が、連続行ではなく、箇条書き等のように 1 行 1 行が別の項目等となっている独立行である場合、ある行の一部をその前後いずれかの行に含めると、文の意味が分かりにくくなり適切ではない。したがって、本実施形態では、文字領域内の行が連続行である場合（S 2 3 で Y E S）と、独立行である場合（S 2 3 で N O）とで、文字画像から文字コードを認識して得られた文字コードデータの配置の仕方が変更される。

## 【 0 0 5 2 】

文字領域内の行が連続行である場合（S 2 3 で Y E S）、1 つの文字領域内に配置することができる文字コードデータ列の幅を計算する（S 2 4）。

## 【 0 0 5 3 】

ここで、注目される文字領域（複数の文字領域のうちの  $i$  番目； $i = 1, 2, \dots$ ）の幅を  $X(i)$ 、行数を  $N(i)$  とすると、1 つの文字領域内に配置することができる文字コードデータ列の幅  $L_{org}(i)$  は、

$$L_{org}(i) = X(i) \cdot N(i)$$

で計算される。

## 【 0 0 5 4 】

一方、実際に配置する文字コードデータ列の幅  $L_{out}(i)$  は、出力に用いるフォントメトリックデータから得られるフォントの文字幅を  $W_{out}(i)$ 、配置

する個々の文字コードデータの数を  $N_c(i)$  とすると、

$$L_{out}(i) = W_{out}(i) \cdot N_c(i)$$

で計算される。

【0055】

なお、各文字コードデータによって文字幅が異なるフォントを用いる場合、それぞれの文字コードデータに対応する文字幅  $W_{out}(c_i)$  を足し合わせる必要があるので、

$$L_{out}(i) = \sum W_{out}(c_i)$$

となる。

【0056】

そして、得られた文字コードデータを出力に用いるフォントで配置した場合、原稿を読み取ることによって入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する (S25)。

【0057】

すなわち、 $L_{out}(i) \leq L_{org}(i)$  ならば、入力された画像データ中の文字領域から文字コードデータがはみだすことはない判別され (S25でYES)、当該文字領域に、文字コードデータを出力フォントで配置する (S31)。

【0058】

一方、 $L_{out}(i) > L_{org}(i)$  ならば、入力された画像データ中の文字領域から文字コードデータがはみだすと判別され (S25でNO)、文字コードデータが配置される文字領域を、原稿を読み取ることによって入力された画像データ中の文字領域に対して、広げることが可能か否かを判別する。すなわち、入力された画像データ中の文字領域に隣接し、かつ、文字画像、または非文字画像である図形画像のいずれをも含まない所定量の未使用部分が存在するか否かを判別する (S26)。

【0059】

図7は、所定量の未使用部分が存在するか否かを判別する方法を説明するための図である。未使用部分が存在するか否かは、入力された画像データ中のある文字領域Aに隣接する左右上下の4方向に、順に一定の幅(長さ)ずつ拡大して調

べていく。

#### 【 0 0 6 0 】

例えば、まず、文字領域 A の右方向に隣接した領域 M1 を調べる。領域 M1 の幅は、出力フォントの文字幅 Wout 程度をとる。ここでは、例えば、領域 M1 の各画素の濃度を多値画像データを用いて調べ、すべての画素の濃度が一定値以下であるときに、領域 M1 は、文字画像、または非文字画像である図形画像のいずれをも含まない未使用部分であると判断する。なお、領域 M1 が未使用部分であるか否かの判断に、2 値画像データを用いることも可能である。

#### 【 0 0 6 1 】

領域 M1 が未使用部分である場合、文字領域を上記未使用部分を含めた領域まで拡大したときに、出力フォントの文字コードデータが拡大後の文字領域からはみだすか否かを再度判別する。例えば領域 M1 が未使用部分でない場合、文字領域 A の右方向には拡大しない。同様にして、領域 M2 ～ M8 について順に未使用部分であるか否かを調べていく。なお、文字領域 A の上下方向に隣接した領域の幅は、出力フォントの文字高さ程度をとる。

#### 【 0 0 6 2 】

このような方法で、文字領域を未使用部分を含めた領域まで順次拡大したときに、出力フォントの文字コードデータが拡大後の文字領域からはみださずに、すべて収まる場合には、所定量の未使用部分が存在すると判断する（S 2 6 で Y E S）。そして、文字領域変更部 3 4 により、文字領域を上記未使用部分を含めた領域まで拡大して、文字コードデータを出力フォントで配置する（S 2 7）。

#### 【 0 0 6 3 】

一方、所定量の未使用部分が存在しなかった場合（S 2 6 で N O）、出力フォントの文字コードデータが文字領域にすべて収まるように、当該文字領域内の文字コードデータを、文字サイズ変更部 3 5 により文字サイズを小さくして配置する（S 3 0）。文字サイズを小さくするために、本実施形態ではフォントサイズを小さくする方法を採用している。但し、当該文字領域内の全体で倍率を小さく指定する方法を採用することも可能である。

#### 【 0 0 6 4 】

なお、文字領域内の行が独立行である場合（S 2 3 で N O）、上記した連続行の場合と同様の処理を行単位で行う。すなわち、文字領域を行単位で分割した個々の行領域内に配置することができる文字コードデータ列の幅を計算する（S 2 9）。また、所定量の未使用部分が存在するか否かは、上記行領域に隣接する左右の 2 方向に、順に一定の幅（長さ）ずつ拡大して調べていく。その他の点は、連続行の場合と同様であるので、詳しい説明を省略する。

## 【 0 0 6 5 】

入力された画像データに複数の文字領域が存在している場合は、すべての文字領域について、上記した出力調整処理（S 1 6、S 2 1～S 3 1）を行う（S 2 8）。

## 【 0 0 6 6 】

出力調整処理が終了するとメインフローチャートに戻り、出力調整処理が行われた文字領域内の文字コードデータは、例えば他の写真領域内の図形画像等とともに、所定の形式のファイルとして出力される（S 1 7）。なお、文字コードデータ等の出力位置は、ファイルの形式によって異なった方法で指定される。

## 【 0 0 6 7 】

このように、本実施形態によれば、入力された画像データ中の文字画像から文字コードを認識して得られた文字コードデータを、ある特定のフォントで配置した場合でも、元のレイアウトを適切に維持しつつ、しかも、文字コードデータを他の文字コードデータや図形画像と重なることなく同一頁内に配置することができる。

## 【 0 0 6 8 】

図 8 は、本発明の画像認識装置の第 2 の実施形態に係るデジタルコピー機を含む情報機器システムの構成を示す概略ブロック図である。以下、第 1 の実施形態との相違点について主に説明する。

## 【 0 0 6 9 】

この実施形態では、図 8 に示すように、画像処理部 1 3 a は、レイアウト解析部 3 1、文字認識部 3 2、フォント認識部 3 3、文字領域変更部 3 4、文字サイズ変更部 3 5、およびファイル作成部 3 6 に加え、文字背景領域検出部 3 7、文



字配置検出部 3 8、および配置形式指定部 3 9 を備えている。

【 0 0 7 0 】

図 9 は、第 2 の実施形態の画像認識処理を示すフローチャート、図 1 0 は、図 9 に示される出力調整処理を示すサブルーチンのフローチャートである。

【 0 0 7 1 】

まず、OCR 前処理 (S 4 1)、レイアウト解析 (S 4 2) が、第 1 の実施形態 (図 2 の S 1 1, S 1 2) と同様に行われた後、文字背景領域の検出が行われる (S 4 3)。

【 0 0 7 2 】

図 1 1 は、原稿を読み取ることにより得られた画像データをレイアウト解析して文字領域、および文字背景領域を検出した結果の一例を示す図である。

【 0 0 7 3 】

図 1 1 に示すように、レイアウト解析部 3 1 により、文字画像が存在する例えば矩形の文字領域 F, G 等が検出されて、抽出される。また、文字背景領域検出部 3 7 により、文字領域 F, G を含み当該文字領域の背景となる背景領域であって、他の異なった背景領域 J の内部に存在する文字背景領域 H, I を検出する。つまり、背景領域 J は、例えば原稿の白地領域に対応しており、文字背景領域 H, I は、例えば原稿の白地領域の内部に存在し、白地領域とは異なった領域として識別される領域に対応している。文字背景領域 H, I を検出するためには、例えば多値画像データで、 $N \times N$  の最小値フィルタ処理 (注目画素の周辺  $N \times N$  画素中の最小値を注目画素の値とする処理) 等を行う。これにより文字画像やラインデータ等を除去した後、均一濃度となっている領域である文字背景領域 H, I を検出することができる。検出結果として、文字背景領域 H, I の位置データと、文字背景領域 H, I 内の色データ (モノクロ画像の場合は濃度データ) が得られる。

【 0 0 7 4 】

次いで、2 値化処理 (S 4 4)、文字コード認識処理 (S 4 5)、フォント認識処理 (S 4 6) が、第 1 の実施形態 (図 2 の S 1 3 ~ S 1 5) と同様に行われた後、出力調整処理が行われる (S 4 7)。

## 【0075】

この出力調整処理では、図10に示すように、まず、行ピッチの算出（S51）、文字領域内の行が連続行であるか否かの判別（S52）が、第1の実施形態（図3のS21，S22）と同様に行われた後、文字領域が文字背景領域に含まれるか否かの判別が行われる（S53）。

## 【0076】

例えば図11に示される文字領域Kのように、文字領域が文字背景領域に含まれない場合（S53でNO）、第1の実施形態（図3のS23～S27，S29～S31）と同様の処理が行われる（S65）。

## 【0077】

一方、文字領域が文字背景領域に含まれる場合（S53でYES）、文字領域内の行が独立行であるときは、文字背景領域内での各行の配置を検出する（S54）。すなわち、文字配置検出部38により、当該各行が、文字背景領域内において左寄せ、右寄せ、またはセンタリングのいずれの形式で配置されているかを検出する。

## 【0078】

図12は、各行が文字背景領域内において左寄せ、右寄せ、またはセンタリングのいずれの形式で配置されているかの判別方法を説明するための図である。この判別は、文字領域内の各行の始点位置、終点位置、中央位置の各x座標についての平均値からのずれの合計を求め、ずれ合計が最も少ないのがいずれであるかで行う。

## 【0079】

すなわち、図12に示すように、 $L_j$ （ $j = 1, 2, \dots$ ），をj行目の始点x座標、 $R_j$ （ $j = 1, 2, \dots$ ）をj行目の終点x座標、 $C_j$ （ $j = 1, 2, \dots$ ）をj行目の中央x座標、それぞれの平均値をL，R，Cとすると、文字背景領域内での文字領域の配置は、

$$\text{Min} \{ \sum (L_j - L), \sum (R_j - R), \sum (C_j - C) \} = \sum (L_j - L)$$

のとき、左寄せ、

$$\text{Min} \{ \sum (L_j - L), \sum (R_j - R), \sum (C_j - C) \} = \sum (R_j - R)$$

のとき、右寄せ、

$$\text{Min} \{ \Sigma (L_j - L), \Sigma (R_j - R), \Sigma (C_j - C) \} = \Sigma (C_j - C)$$

のとき、センタリング、

と判別することができる。

#### 【 0 0 8 0 】

また、文字領域内の行が連続行である場合（S 5 5 で Y E S）、第 1 の実施形態（図 3 の S 2 4）と同様にして、1 つの文字領域内に配置することができる文字コードデータ列の幅を計算する（S 5 6）。一方、文字領域内の行が独立行である場合（S 5 5 で N O）、第 1 の実施形態（図 3 の S 2 9）と同様にして、文字領域を行単位で分割した個々の行領域内に配置することができる文字コードデータ列の幅を計算する（S 6 2）。

#### 【 0 0 8 1 】

そして、得られた文字コードデータを出力に用いるフォントで配置した場合、原稿を読み取ることによって入力された画像データ中の文字領域から文字コードデータがはみだすか否かを判別する（S 5 7）。入力された画像データ中の文字領域から文字コードデータがはみださない場合は（S 5 7 で Y E S）、上記第 1 の実施形態と同様に（図 3 の S 3 1）、当該文字領域に、文字コードデータを出力フォントで配置する（S 6 4）。

#### 【 0 0 8 2 】

但し、この第 2 の実施形態では、入力された画像データ中の文字領域から文字コードデータがはみだす場合（S 5 7 で N O）の処理が第 1 の実施形態と相違している。すなわち、第 1 の実施形態のように左右上下方向に隣接する所定量の未使用部分が存在するか否かを判別するのではなく、文字コードデータが配置される文字領域を、文字背景領域の範囲内で左右または上下方向に均等に、文字コードデータが収まるまで拡大していったときに、文字コードデータが文字背景領域の範囲内で収まるか否かを判断する（S 5 8）。

#### 【 0 0 8 3 】

文字コードデータが文字背景領域の範囲内で収まる場合（S 5 8 で Y E S）、文字領域変更部 3 4 により、文字コードデータが配置される文字領域を文字背景

領域の範囲内で拡大して、文字コードデータを出力フォントで配置する（S 5 9）。

#### 【 0 0 8 4 】

一方、文字コードデータが配置される文字領域を文字背景領域と一致するまで拡大した場合でも、文字コードデータが文字背景領域の範囲内で収まらない場合（S 5 8 で N O）、出力フォントの文字コードデータが文字領域にすべて収まるように、当該文字領域内の文字コードデータを、文字サイズ変更部 3 5 により文字サイズを小さくして配置する（S 6 3）。

#### 【 0 0 8 5 】

次に、文字領域内の行が独立行であるときは、文字配置検出部 3 8 の検出結果に基づいて、配置形式指定部 3 9 により、各行の配置形式を指定する（S 6 0）。

#### 【 0 0 8 6 】

入力された画像データに複数の文字領域が存在している場合は、すべての文字領域について、上記した出力調整処理（S 4 7、S 5 1 ～ S 6 5）を行う（S 6 1）。

#### 【 0 0 8 7 】

出力調整処理が終了するとメインフローチャートに戻り、出力調整処理が行われた文字領域内の文字コードデータは、例えば他の写真領域内の図形画像等とともに、所定の形式のファイルとして出力される（S 4 8）。なお、文字背景領域内における左寄せ、右寄せ、またはセンタリングの配置形式を含む文字コードデータ等の出力位置は、ファイルの形式によって異なった方法で指定される。

#### 【 0 0 8 8 】

このように、第 2 の実施形態によれば、第 1 の実施形態と同一の効果を得ることができることに加え、文字領域が文字背景領域に含まれている場合には、得られた文字コードデータを、ある特定のフォントで配置した場合でも、文字背景領域の範囲内で配置することができ、元のレイアウトを適切に維持しつつ、しかも、文字コードデータを他の文字コードデータや図形画像と重なることなく同一頁内に配置することができる。また、文字背景領域内における各行の左寄せ、右寄

せ、またはセンタリングの形式を保って配置させることができる。

【 0 0 8 9 】

本発明は、上記した実施形態のみに限定されるものではなく、特許請求の範囲内において、種々改変することができる。

【 0 0 9 0 】

例えば、上記実施形態では、画像認識装置としてデジタルコピー機を例に挙げて説明したが、本発明の画像認識装置はこれに限定されるものではなく、例えばスキャナから画像データを受信して同様の処理をコンピュータで行うシステム、同様の処理をスキャナで行ってコンピュータやプリンタ等に送信するシステム、スキャナから直接画像データを受信してプリンタで同様の処理を行うシステム等にも適用することが可能である。

【 0 0 9 1 】

なお、上記した実施形態において、画像認識の制御は、上記した処理手順（図 2、図 3 等参照）を記述した所定のプログラムを主制御部 1 1 が実行することによって行われるものであり、この所定のプログラムは、コンピュータ読取可能な記録媒体（例えば、フロッピーディスクや CD-ROM 等）によって提供されることもできる。また、この所定のプログラムは、上記各処理を実行するアプリケーションソフトウェアとして提供されてもよいし、コンピュータやデジタルコピー機等の各情報機器や管理サーバの一機能として各情報機器や管理サーバのソフトウェアに組み込んでもよい。

【 0 0 9 2 】

【発明の効果】

以上説明したように、本発明によれば、入力された画像データ中の文字画像から文字コードを認識して得られた文字コードデータを、ある特定のフォントで配置した場合でも、元のレイアウトを適切に維持しつつ、しかも、文字コードデータを他の文字コードデータや図形画像と重なることなく同一頁内に配置することができる。

【 0 0 9 3 】

また、文字領域が文字背景領域に含まれている場合、文字背景領域を検出する

ことにより、得られた文字コードデータを、ある特定のフォントで配置した場合でも、文字背景領域の範囲内で配置することができる。さらに、文字背景領域内で文字画像が複数の独立した列を形成している場合、各列が文字背景領域内において左寄せ、右寄せ、またはセンタリングのいずれの形式で配置されているかを検出することにより、当該配置形式を保つことができる。

【図面の簡単な説明】

【図 1】 本発明の画像認識装置の第 1 の実施形態に係るデジタルコピー機を含む情報機器システムの構成を示す概略ブロック図である。

【図 2】 画像認識処理を示すフローチャートである。

【図 3】 図 2 に示される出力調整処理を示すサブルーチンのフローチャートである。

【図 4】 原稿の一例を示す図である。

【図 5】 図 3 の原稿を読み取ることによって入力された画像データをレイアウト解析して文字領域を検出した結果を示す図である。

【図 6】 文字領域内の行が連続行であるか否かの判別方法を説明するための図である。

【図 7】 所定量の未使用部分が存在するか否かを判別する方法を説明するための図である。

【図 8】 本発明の画像認識装置の第 2 の実施形態に係るデジタルコピー機を含む情報機器システムの構成を示す概略ブロック図である。

【図 9】 第 2 の実施形態の画像認識処理を示すフローチャートである。

【図 10】 図 9 に示される出力調整処理を示すサブルーチンのフローチャートである。

【図 11】 原稿を読み取ることにより得られた画像データをレイアウト解析して文字領域、および文字背景領域を検出した結果の一例を示す図である。

【図 12】 各行が文字背景領域内において左寄せ、右寄せ、またはセンタリングのいずれの形式で配置されているかの判別方法を説明するための図である。

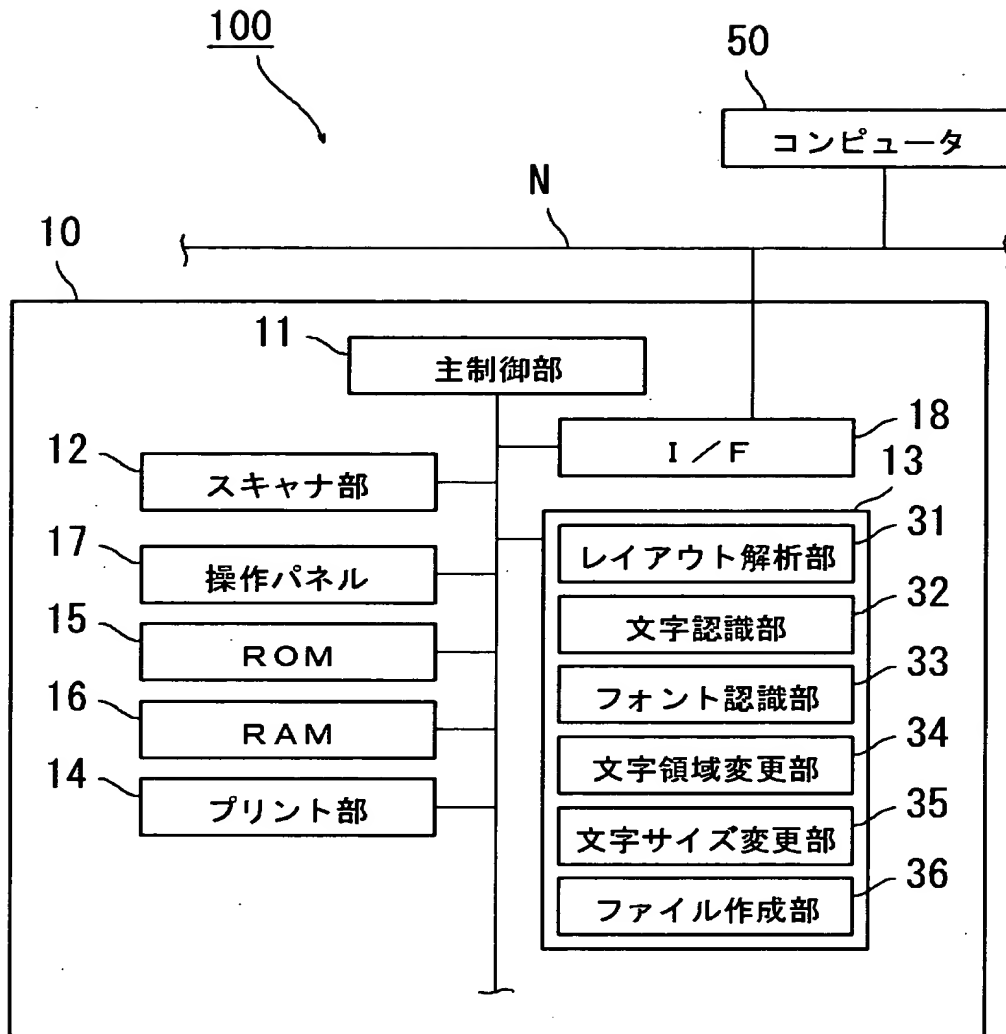
【符号の説明】

- 1 0 …デジタルコピー機（画像認識装置）、
- 3 1 …レイアウト解析部（文字領域検出部）、
- 3 2 …文字認識部、
- 3 4 …文字領域変更部、
- 3 5 …文字サイズ変更部、
- 3 7 …文字背景領域検出部、
- 3 8 …文字配置検出部、
- 3 9 …配置形式指定部、
- A ～ G …文字領域、
- H, I …文字背景領域。

【書類名】

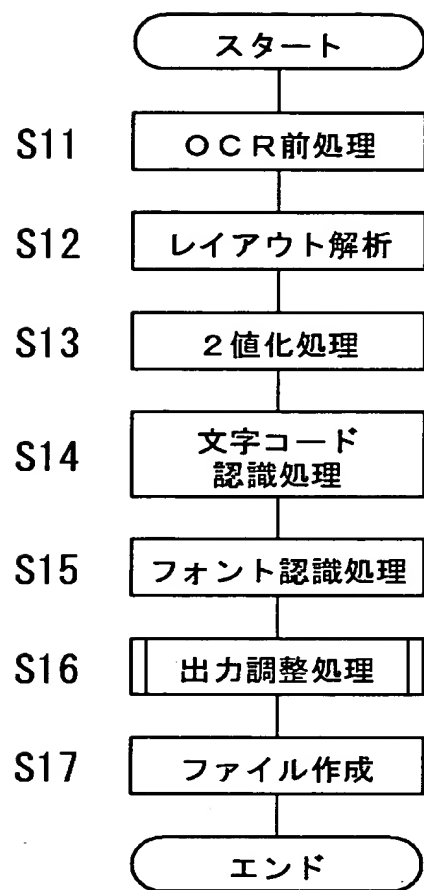
図面

【図 1】

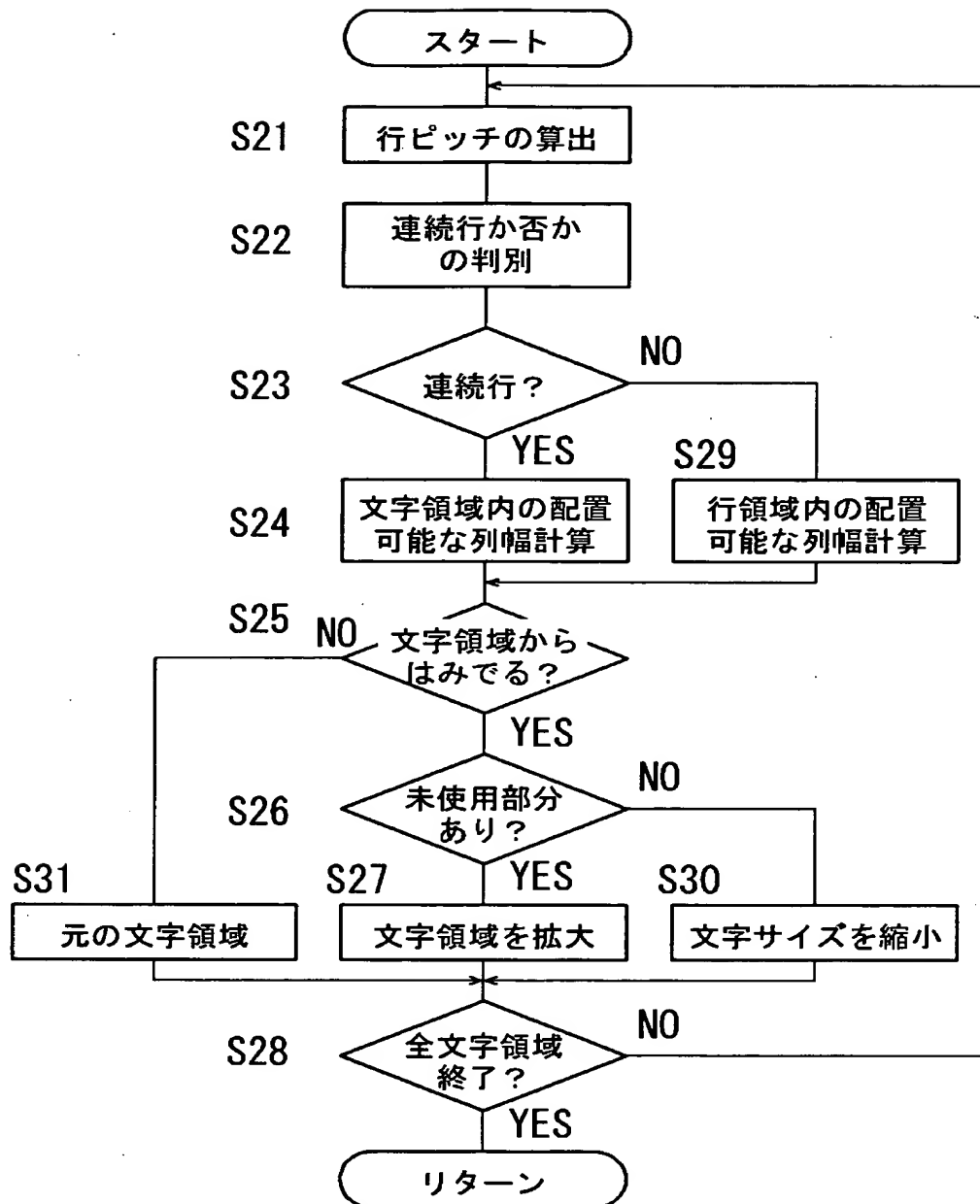




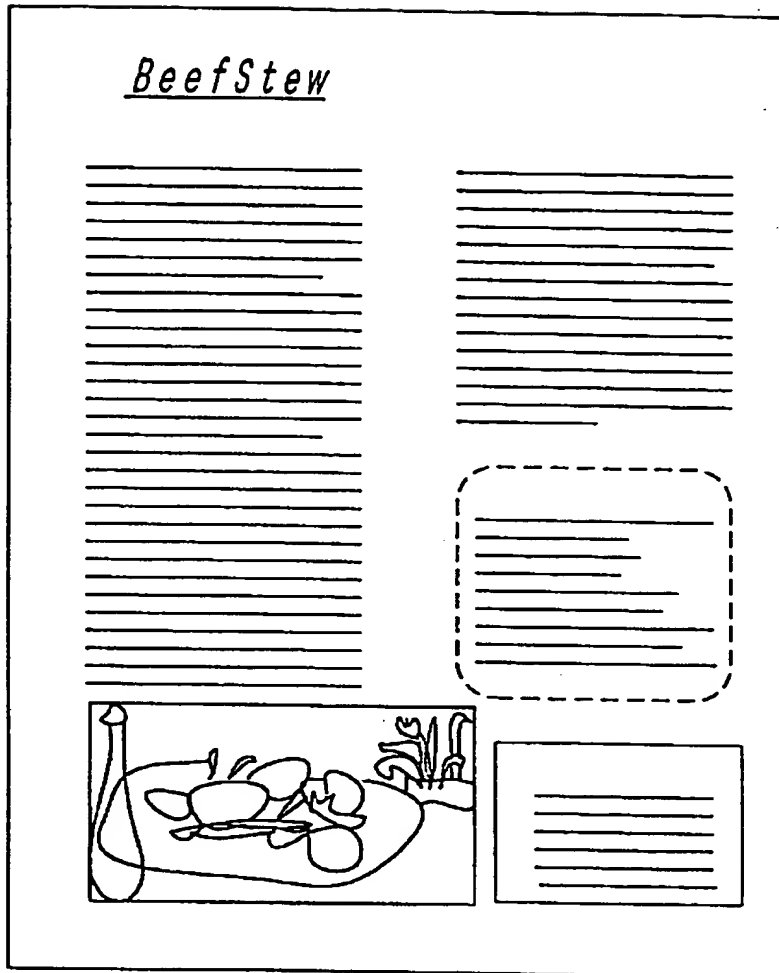
【図 2】



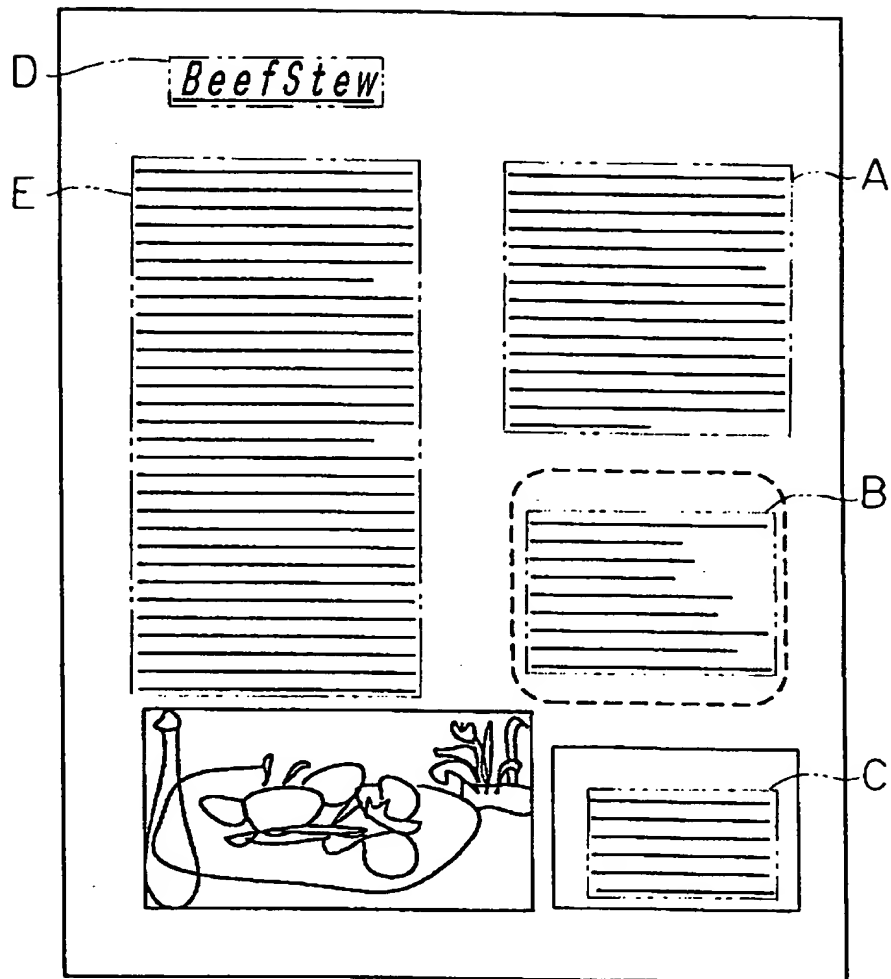
【図 3】



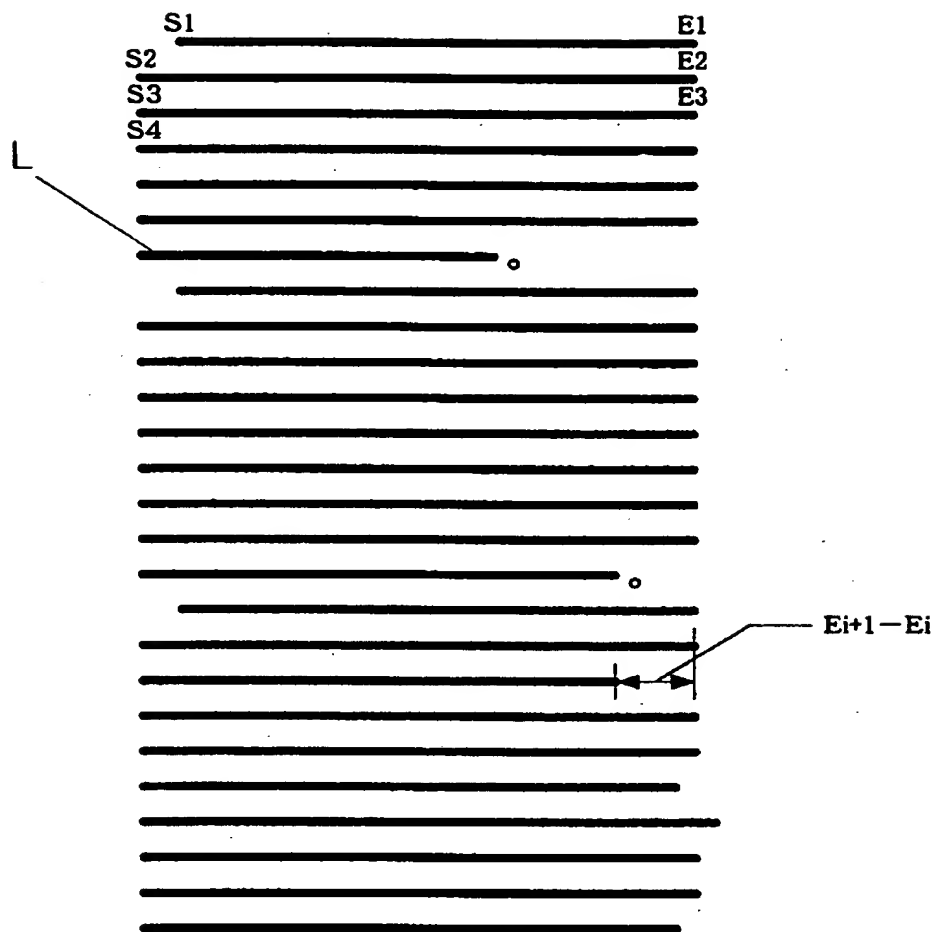
【図 4】



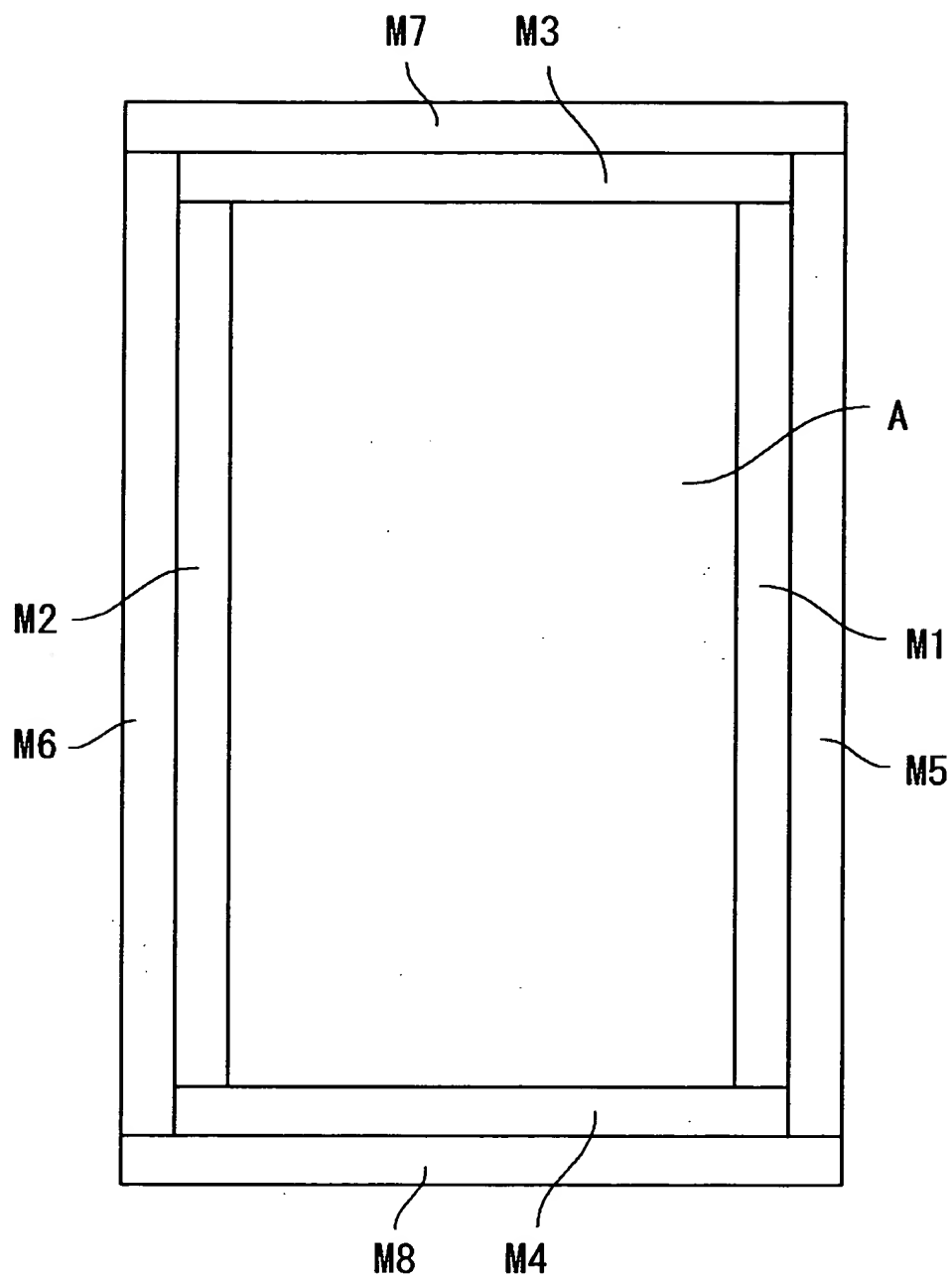
【図 5】



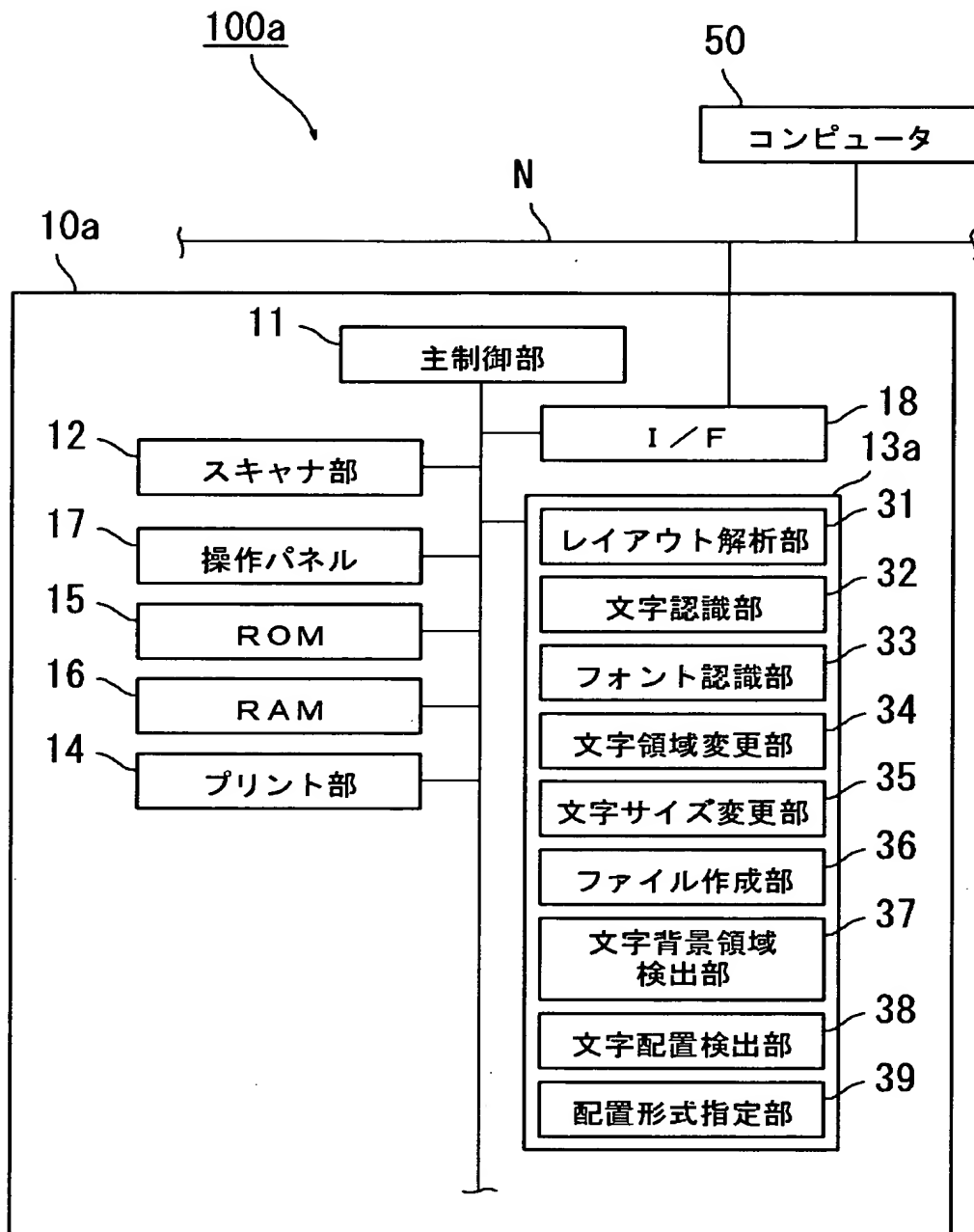
【図 6】



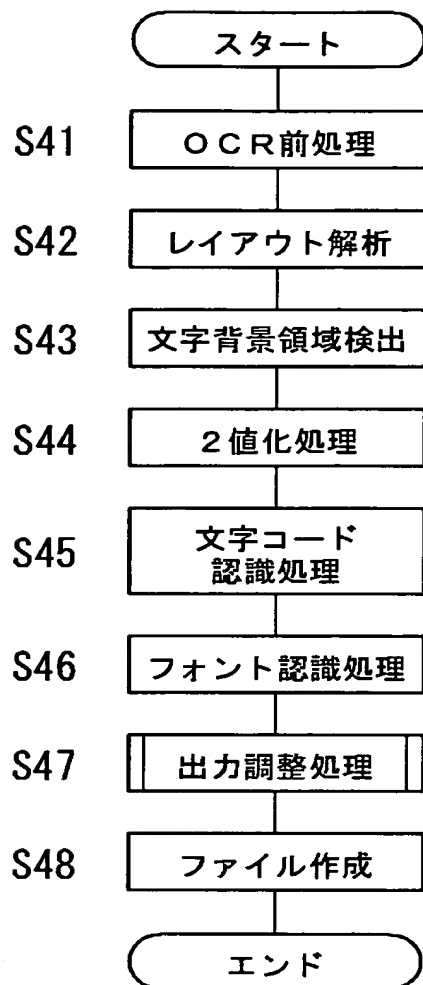
【図 7】



【図 8】

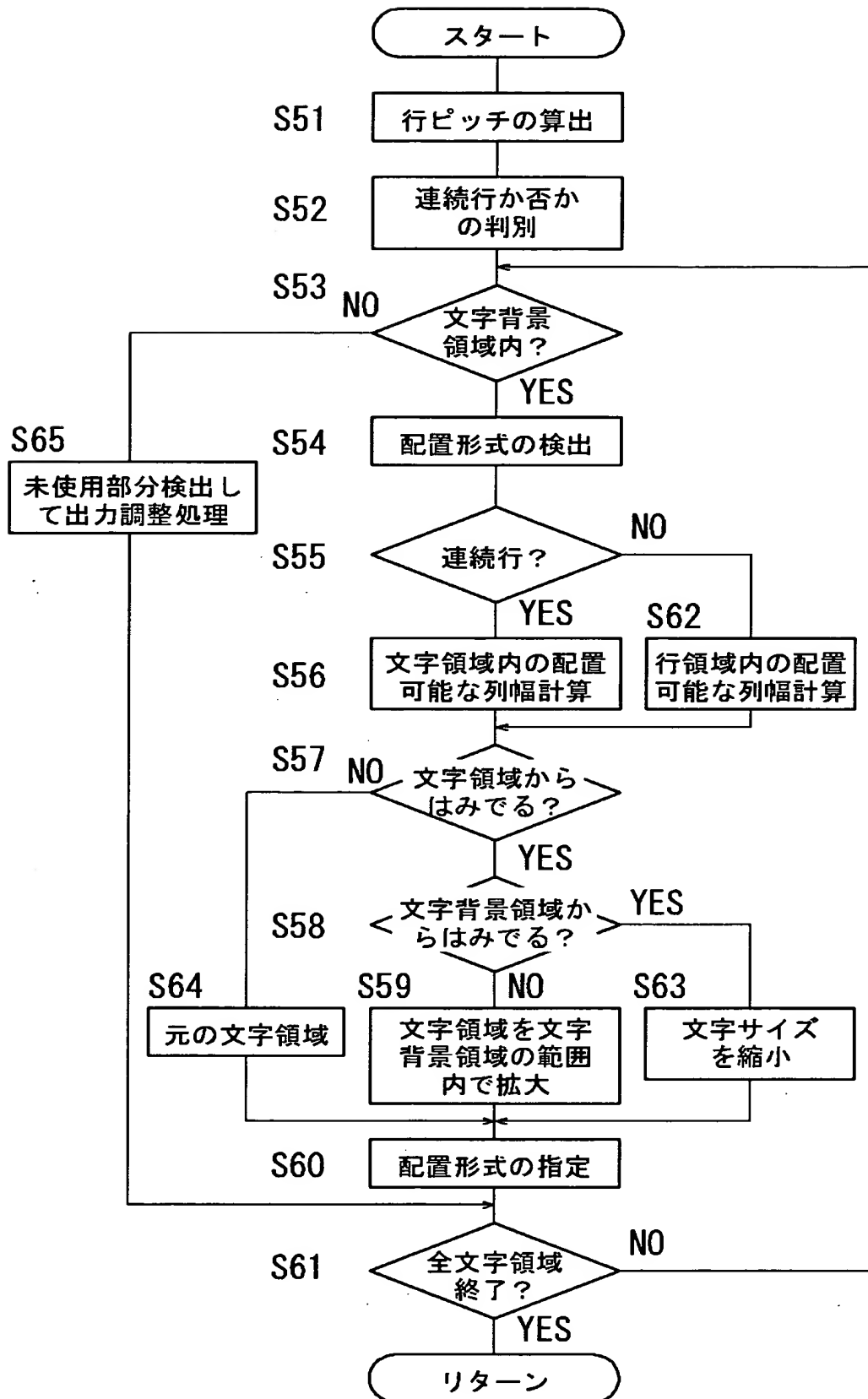


【図 9】

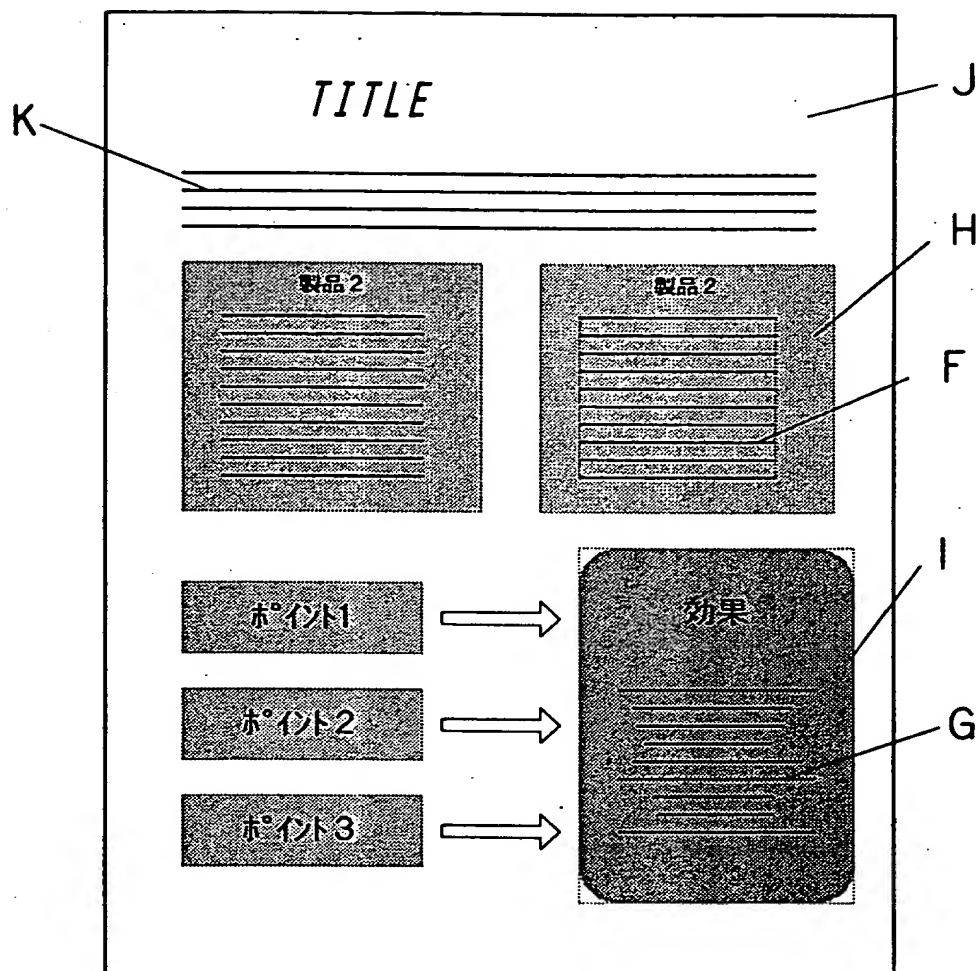




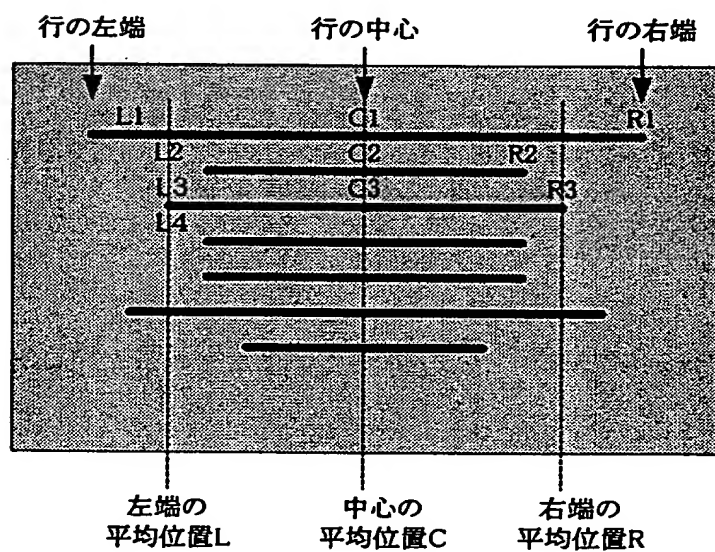
【図 1 0】



【図 1 1】



【図 1 2】



【書類名】                      要約書

【要約】

【課題】    入力された画像データ中の文字画像から文字コードを認識して得られた文字コードデータがある特定のフォントで配置した場合でも、元のレイアウトを適切に維持しつつ、しかも、文字コードデータを他の文字コードデータや図形画像と重なることなく同一頁内に配置する。

【解決手段】    画像認識装置としてのデジタルコピー機 1 0 は、入力された画像データ中の文字領域から、文字コードを認識して得られた文字コードデータがはみだすと判別された場合であって、文字領域に隣接し、かつ、文字画像、または非文字である図形画像のいずれをも含まない所定量の未使用部分が存在するとき、文字領域変更部 3 4 により、文字コードデータが配置される文字領域を、上記未使用部分を含めた領域まで拡大する。

【選択図】                      図 1

出 願 人 履 歴 情 報

識別番号 [ 0 0 0 0 0 6 0 7 9 ]

1. 変更年月日 1 9 9 4 年 7 月 2 0 日

[変更理由] 名称変更

住 所 大阪府大阪市中央区安土町二丁目 3 番 1 3 号 大阪国際ビル  
氏 名 ミノルタ株式会社